

## GROUPING METHOD OF IMAGE FRAGMENTS OF ADJACENT DISLOCATION ETCH PITS OF THE SEMICONDUCTOR WAFER

*Andrey Samoilov*

*Department of Information and Control Systems  
Kremenchuk Mykhailo Ostrohradskyi National University  
20 Pershotravneva str., Kremenchuk, Ukraine, 39600  
seticom@yandex.ru*

*Igor Shevchenko*

*Department of Information and Control Systems  
Kremenchuk Mykhailo Ostrohradskyi National University  
20 Pershotravneva str., Kremenchuk, Ukraine, 39600  
silver055@mail.ru*

---

### Abstract

An increase in production volumes of gallium arsenide semiconductor devices determines the need for better control of dislocations of semiconductor wafer.

The grouping method of image fragments of adjacent dislocation etch pits of the semiconductor wafer is proposed in the article. Adjacent fragments will be allocated in the pre-binarized image of wafer surface, which contains adjacent fragments of etch pits of dislocation loops after treatment by the described method. Improved methods for determining the loop line width determines the edge line width of etch pits of suspected dislocations, given the variability of their display in the binarized image. The current loop line width is compared to the reference line width of the dislocation loop.

The grouping method of image fragments of adjacent dislocation etch pits of the semiconductor wafer defines recovery of loop lines branching, takes into account various options of line adjacency and determines the direction of further recovery of loop line of dislocation etch pits. A step by step description of the method is given.

**Keywords:** etch pits, dislocation, loop fragments, gallium arsenide, digital image.

DOI: 10.21303/2461-4262.2016.00128

© Andrey Samoilov, Igor Shevchenko

### 1. Introduction

In the process of growth of gallium arsenide single crystal, especially using the Czochralski method, the dislocation can be formed as a single and on the adjacent edges.

The use of selective etching and light microscopy to control the amount of dislocation density give different contrasting reflectance for the etch pits in the semiconductor wafer plane. Changes in the intensity of reflected light are not always the same.

In addition, there are other contrast pieces, reflecting a structural imperfection that does not belong to the category of dislocation etch pits in the semiconductor wafer. Therefore, counting the dislocation density as a significant indicator of structural imperfection is very time-consuming. It is especially difficult to take into account the adjacent dislocation. The increased dislocation density reduces the quality of the product and narrows its scope.

Thus, given the complexity of calculating the adjacent dislocations and importance of dislocation density index, the task of grouping of adjacent fragments of dislocation etch pits loop is relevant.

Diversity of displaying dislocation etch pits provided in a contrasting image of semiconductor wafer surfaces [1] create some limitations for the use of clustering methods to obtain information about the adjacency degree of binarized fragments of dislocation etch pits. Given this fact, it should consider clustering methods as a prospective tool to get information about the clusters of adjacent fragments in the image of semiconductor wafer surface.

Hierarchical and non-hierarchical clustering methods are considered in [2]. Hierarchical methods show the results in the form of reachability diagram, dendrogram or mosaic view.

There are disadvantages of hierarchical methods for grouping of adjacent fragments of etch pits:

- umbersomeness on large data sets, and decrease has negative impact on the results;
- high computational complexity, including for dissimilar loop fragments of etch pits;
- critical to choosing proximity measure;
- redundancy of full partition system for the problem of grouping of adjacent fragments of etch pits.

Non-hierarchical methods can be carried out processing in the image with less quality, containing noise and non-uniformity leading to data emissions, the appearance of minor data in data array. Presentation of the results is possible as a linear projection, nomogram etc. However, before performing the clustering it should be known in advance the number of clusters, iterations and stopping criteria for calculation.

There are the most well-known non-hierarchical methods in terms of applicability for grouping of adjacent fragments of etch pits:

- fuzzy clustering [2, 3];
- partitioning [4];
- artificial neural networks for clustering [5];
- grid-based [6];
- density-based [7].

Density-based clustering methods [7] are divided into two types. First: density-based connectivity methods. Typical methods are DBSCAN [8], OPTICS [9]. Found connectivity to all points accessible from the main objects can be reflected in the maximum connected components, implying clusters. Points that are not associated with any reference point don't fall within the cluster.

The second type of density-based clustering methods uses calculation of density functions. A typical method is DENCLUE [10]. The method focuses not on the calculated data density, and on the density function for the attribute space.

Low level of DBSCAN formalization and sensitivity to the choice of parameters deprive efficiency and automation of operation. Adjacent groups (clusters) have complex shapes and sharp changes in the number of fragments taken into account in the cluster do not allow them to be connected together.

OPTICS method [9] is an advanced DBSCAN method. OPTICS algorithm uses one parameter for selection of optimal solution. These clusters take into account changes in the density of objects and can be connected, eliminating fragmentation of adjacent clusters.

The disadvantages of OPTICS method are: relatively slow (about in 1.6 times) data processing in comparison with DBSCAN [8] and non-operational, non-automatic selection of parameter  $\epsilon$  (maximum radius for neighbor accounting).

DENCLUE method is described in [10]. This method takes into account the disadvantages and synthesizes OPTICS and DBSCAN methods on the basis of accounting density of fragments.

Function of Parzen density estimation in DENCLUE method is effectively reduces the effect of noise. Local maxima are found using “mean shift” procedure proposed in [8] and was first used for image segmentation in [10]. The “mean shift” procedure uses the points in dense cells as a source cell and neighboring cells. Therefore DENCLUE well works with noise and abrupt changes in the data. Hinneburg and Gabriel [11] developed DENCLUE 2.0 method, which uses a new procedure of increase for Gaussian kernels, automatically adjusting the step size.

Advantages of DENCLUE: invariance to noise due to kernel density estimation by the uniform distribution of noise in the input data for areas with very dense arrangement of the objects.

Disadvantages of DENCLUE method are: large computational efforts of grid “mean shift” procedure, clusters allocated by “mean shift” procedure are characterized by fragmentation, which is contrary to the task of grouping of fragments on the semiconductor wafer.

The computational complexity of DBSCAN OPTICS and DENCLUE methods is the same and defined as  $O(N \log N)$ , where  $N$  – number of grouping fragments.

Overall, disadvantages of presented density-based methods are:

- the need for clear guidance of the number of clusters and points involved in it (DBSCAN);
- density estimation are sensitive to the radius of considered neighborhood (DBSCAN and OPTICS);
- “mean shift” procedure of DENCLUE method is time consuming for clustering images of the semiconductor wafers, as well as the fragmentation of the clusters. which contradicts the task of grouping of fragments.

Thus, given the disadvantages of using DBSCAN, OPTICS and DENCLUE clustering methods possessing the above characteristics, these density-based methods are not able to reflect the specificity of task for searching adjacent compact groups of fragments. To solve the problem of grouping of fragments of dislocation etch pits it is necessary to use specialized grouping method of image fragments of adjacent dislocation etch pits of the semiconductor wafer.

## 2. Materials and methods

First of all, developed method must has following properties:

- the minimum number of adjustable parameters;
- automatic operation;
- ability to study fragments reflecting the adjacent groups of arbitrary shape;
- scalability and ability to work with data of fragments for semiconductor wafers without adjustment;
- ability to process large amounts of data with the acceptable time;
- ability to detect and eliminate possible deviations of grouping of fragments from adjacent groups;
- simple settings;
- possibility of obtaining information on the number of potential clusters without pre-processing;
- providing results to facilitate further analysis.

Formation of etch pit loop from the fragments using the methods of loop restore fails due to the location adjacency of the main significant fragments. It is logical to assume that the increased density of the location of the main significant loop fragments in the plane of binarized image may indicate a group of adjacent and possibly several randomly overlapping fragments of some dislocation etch pits.

It is proposed to use such change in the density of the main significant fragments of dislocation etch pits as a sign of their accommodation.

All fragments in the plane of a digital raster image can be divided into significant and insignificant fragments [12].

A plurality of fragments A for binarized image with accentuated contours after applying adaptive binarization J2 with parameter C=15 [12, 13] can be described by the formula:

$$A = (B \cup C), \quad (1)$$

where A – the set of fragments of binarized image of the semiconductor wafer; B – a significant subset – part of the fragments defining basic parts of loop lines of etch pits in the image; C – the remaining fragments of set A are not included in the subset B.

Adjacency C with meaningful B fragments suggests that they belong to particular region of considered dislocation etch pits into a digital raster image plane.

A subset of significant fragments may include closed or open single loops satisfying loop line width [14].

Insignificant fragments – other fragments available on the binarized image of the semiconductor wafer with accentuated contours J2.

In the case of search adjacent fragments of dislocation etch pits loops there are no need for the separation of fragments according the classes determined by the number and mutual arrangement.

Adjacency grouping of fragments that available in the image simplifies the task of clustering by reducing the number of required classes to two: adjacent and non-adjacent fragments of the etch pits of B – significant subset, i. e. fragments defining the main parts of etch pits loops in the image.

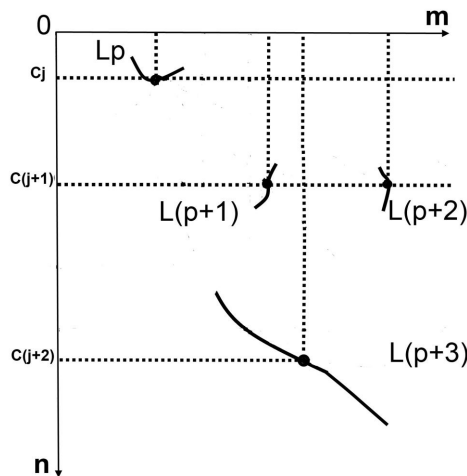
The studied plane of binarized image has fragments of dislocation etch pits of semiconductor wafer, which describes the intensity  $I = 255$  and  $I = 0$ . Fragments for grouping are the intensity  $I = 0$  and the rest surface  $I = 255$ .

The wafer plane in a digital raster image is pre-represented in the Cartesian coordinate system with the dimension of  $m \times n$  pixels, where  $m$  – the dimension of the raster in image pixels on the abscissa axis, and  $n$  – the dimension of the raster in image pixels on the ordinate (**Fig. 1**).

Finding the nearest center of mass of the fragment in the binary image with intensity  $I = 0$  using the test line  $C_n$ .

Transition grid step to the coordinates has a pixel dimension for the binarized image with accentuated loops after applying adaptive binarization J2 to a set of fragments A. However, the transition itself is possible to advance the known centers of mass of fragments B of meaningful subsets, which speeds up the grouping method.

The direction of research is carried out in the general order of increasing coordinate axes  $m$  and  $n$ .



**Fig. 1.** Total centers of mass processing circuit of dislocation fragments in the plane of the digital raster image of GaAs wafer represented as coordinate axes  $m$ ,  $n$  on the basis of determining the intersection of test lines  $C_j$ ,  $C_{(j+1)}$ ,  $C_{(j+2)}$  with the fragments  $L_p$ ,  $L_{(p+1)}$ ,  $L_{(p+2)}$ ,  $L_{(p+3)}$  in coordinates of the centers of mass  $X_{Mp}$ ,  $Y_{Mp}$ ,  $X_{M(p+1)}$ ,  $Y_{M(p+1)}$ ,  $X_{M(p+2)}$ ,  $Y_{M(p+2)}$ ,  $X_{M(p+3)}$  with intersection points that found in the coordinates respectively. It is shown not to scale

The number of operations in this case increases in proportion to the number of desired fragments for grouping in the image J2.

Increment step  $m$  along the abscissa axis of the current coordinates  $x_i$  is equal to unit pixel of the image, where  $i = (0, \dots, m)$ . Unit increment step on the ordinate to the coordinates  $y_j$  where  $j = (0, \dots, n)$ .

The test lines  $C_j$  examined the plane of binarized image (**Fig. 1**) with a minimum increment equal to one point of the image of corresponding coordinates. However, given the known coordinates of the center of mass  $X_{Mp}$  and  $Y_{Mp}$ , increment step for desired fragments can be increased consistent with the current value of the center of mass to accelerate the grouping.

Test line  $C_j$  with the ordinate  $y_{jp}$  for the fragment  $L_p$  gets the number  $p$ , where  $p = (0, \dots, F)$ , where  $F$  – the maximum number of fragments  $L_F$  in the image J2.

The form of presentation of objects to developed grouping method is sign of center of mass  $X_{Mp}$  and  $Y_{Mp}$  of each of the considered fragments in the coordinate of axes  $m$  and  $n$  by determining intersections of the test lines  $C_j$  with fragments of etch pits  $L_p$ .

Selection of proximity measure effects on the clustering results that may differ significantly when using different measures.

To developing grouping methods of image fragments are encouraged to use the maximum distance between the sides of alleged dislocation etch pit Z as a proximity measure.

This characteristic is determined by the standard dislocation etch pit [13].

Proximity measure determines the adjacency of fragments (**Fig. 2**). Adjacency evaluation of two points A and C is also occurs on the basis of determining the proximity measure Z between the centers of mass on plane with coordinates  $X_{M_A}$ ,  $Y_{M_A}$  and  $X_{M_C}$ ,  $Y_{M_C}$ . The proximity measure is calculated as the hypotenuse of a right triangle formed by the coordinates centers of mass  $X_{M_A}$ ,  $Y_{M_A}$  and  $X_{M_C}$ ,  $Y_{M_C}$ , points E, A and B (**Fig. 2**). The following formula determines the inequality of belonging to a group of adjacent fragments of developed grouping method:

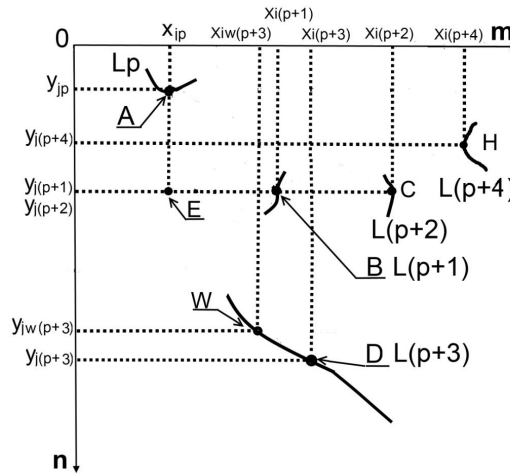
$$\sqrt{(Y_{M_A} - Y_{M_C})^2 + (X_{M_A} - X_{M_C})^2} \leq Z, \quad (2)$$

where the proximity of the fragments is calculated in the left-hand side and reflects the hypotenuse calculation between points A and C on the legs determined by the difference between the absolute coordinates of the points:  $X_{M_A} - X_{M_C}$ ,  $Y_{M_A} - Y_{M_C}$  for the abscissa and ordinate axes respectively.

For fragment  $L_{(p+1)}$  (**Fig. 2**) of the semiconductor wafer of point B the inequality of belonging to grouping of adjacent fragments will be:

$$\sqrt{(Y_{M_p} - Y_{M_{(p+1)}})^2 + (X_{M_p} - X_{M_{(p+1)}})^2} \leq Z, \quad (3)$$

where the proximity of the fragments is calculated on the left-hand side and reflects the hypotenuse calculation on the legs determined by the difference between the absolute coordinates of the points  $X_{M_p} - X_{M_{(p+1)}}$ ,  $Y_{M_p} - Y_{M_{(p+1)}}$  for the abscissa and ordinate axes respectively.



**Fig. 2.** The detection circuit of adjacent loops: at the point A of the fragment  $L_p$  with the current coordinates  $x_{ip}$ ,  $y_{jp}$ , at the point B of the fragment  $L_{(p+1)}$   $x_{i(p+1)}$ , at the point C of the fragment  $L_{(p+2)}$  with the current coordinates  $x_{i(p+2)}$ ,  $y_{i(p+2)}$ , partially adjacent fragment  $L_{(p+3)}$  at the point D  $x_{i(p+3)}$ ,  $y_{i(p+3)}$  by detecting a point W with coordinates  $x_{iw(p+3)}$ ,  $y_{iw(p+3)}$  and a fragment  $L_{(p+4)}$  at the H point with the current coordinates  $x_{i(p+4)}$ ,  $y_{i(p+4)}$ . It is shown not to scale

Successful search of adjacent to the  $L_p$  fragment  $L_{(p+1)}$  with the coordinates of the center of mass  $X_{M_{(p+1)}}$  and  $Y_{M_{(p+1)}}$  occurs if  $x_{i(p+1)} = X_{M_{(p+1)}}$  and  $y_{i(p+1)} = Y_{M_{(p+1)}}$  and condition (3) is performed for points A and B.

Search of fragments on abscissa occurs until  $i = m$  consider the example of the current coordinate  $x_i$  using data about nearest centers of mass  $X_{M_{(p+2)}}$  at the point C in **Fig. 2**.

Detection of adjacent to the  $L_{(p+1)}$  fragment  $L_{(p+2)}$  at the point C with coordinates of center of mass  $X_{M_{(p+2)}}$  and  $Y_{M_{(p+2)}}$  occurs if the condition (3) is performed within the radius Z. Successful search condition of adjacent fragment  $L_{(p+2)}$  is found fragment in the coordinates  $x_{i_{(p+2)}} = X_{M_{(p+2)}}$  and  $y_{i_{(p+2)}} = Y_{M_{(p+2)}}$ .

Grouping is also done for fragments with the centers of mass that removed more than a proximity measure distance (3), but have some parts that fall within the radius Z.

Partially adjacent fragment  $L_{(p+3)}$  is grouped by the point W, satisfying proximity (3) within the radius Z, in the current coordinates  $x_i = x_{iw_{(p+3)}}$ ,  $y_j = y_{iw_{(p+3)}}$ .

Directly grouping of fragment  $L_{(p+3)}$  is performed in the coordinates of the center of mass  $X_{M_{(p+3)}}$  and  $Y_{M_{(p+3)}}$ , the search of which is realized by the loop recovery to achieve coordinates  $x_{i_{(p+3)}}$ ,  $y_{i_{(p+3)}}$ . The recovery method is described in [14].

If adjacent fragments are not detected, a transition of current coordinates to the next coordinate of the center of mass along the ordinate axis.

If the detection occurs again for any centers of mass, the second time such fragments are ignored.

After achieving  $y_{i_p}$  by the current coordinates on the ordinate axis of the limiting point  $j = n$  it is necessary to continue the search for fragments along the abscissa axis from the center of coordinate on the ordinate axis. The condition for a successful search of the adjacent fragment  $L_{(p+4)}$  (Fig. 2) is the following found point of H fragment  $L_{(p+4)}$  in coordinates  $x_{i_{(p+4)}} = X_{M_{(p+4)}}$  and  $y_{i_{(p+4)}} = Y_{M_{(p+4)}}$ .

Thus, after adjacency analysis of all significant fragments will obtain two classes of fragments: adjacent and grouped relating to etch pits, and other containing significant and insignificant fragments.

### 3. Results of research

The proposed grouping method of image fragments of adjacent dislocation etch pits of the semiconductor wafer is a development of the nearest neighbor method. It allows to form groups of fragments of dislocation etch pits, allowing calculation of the number of adjacent dislocations on the surface of the semiconductor wafer after the selective etching of the semiconductor wafer.

The developed grouping method of image fragments of adjacent dislocation etch pits of the semiconductor wafer uses such input data:

- attributive description – fragment is described by numerical attribute of line thickness, a sign of the center of mass with coordinates  $X_{Mp}$  и  $Y_{Mp}$ ;
- distance between the fragments. During the grouping of adjacent fragments the maximal distance between the centers of mass of etch pits loop of reference dislocation Z is used.

Grouping method of image fragments of adjacent dislocation etch pits of the semiconductor wafer has the following steps:

1. Feature array of centers of mass  $X_{Mp}$  и  $Y_{Mp}$  of all relevant intensity fragments  $I = 0$  sorted by increase of ordinate values is used for the image J2 in pixel  $m \times n$  coordinate system, where  $m$  – the dimension of the raster in pixels of image on the abscissa axis and  $n$  – the dimension of the raster in pixels of image on the ordinate axis (Fig. 1).

2. Transition through the value array of centers of mass of fragments until the first unexplored fragment, the first time it is a fragment  $L_p$  using the current coordinates  $x_i$  to the point  $x_{i_p}$  (Fig. 2). Detection of the nearest center of mass  $X_{Mp}$   $Y_{Mp}$  of the fragment  $L_p$  on the condition of equality of current coordinates and coordinates of the center of mass of fragment  $x_i = X_{Mp}$  and  $y_{j_p} = Y_{Mp}$ .

3. Detection of the adjacent fragment  $L_{(p+1)}$  at the point B with coordinates of the center of mass  $X_{M_{(p+1)}}$  and  $Y_{M_{(p+1)}}$ , if  $x_{i_{(p+1)}} = X_{M_{(p+1)}}$  and  $y_{i_{(p+1)}} = Y_{M_{(p+1)}}$  and the condition of belonging to a grouping of adjacent fragments (3) is performed for points A and B.

4. Detection of adjacent fragment  $L_{(p+2)}$  at the point C with coordinates of the center of mass  $X_{M_{(p+2)}}$  and  $Y_{M_{(p+2)}}$ , and if  $x_{i_{(p+2)}} = X_{M_{(p+2)}}$  and  $y_{i_{(p+2)}} = Y_{M_{(p+2)}}$ . The condition of belonging to a grouping of adjacent fragments (3) is determined from point B to detected point C.

5. Detection of partially adjacent fragment (Fig. 2) after detecting point W in the current coordinates  $x_{iw_{(p+3)}}$ ,  $y_{iw_{(p+3)}}$  on the condition of belonging to a grouping of adjacent fragments (3). De-



tection of the center of mass  $x_{i_{(p+3)}}$ ,  $y_{i_{(p+3)}}$  of partially adjacent fragment  $L_{(p+3)}$  by the loop recovery method [14], if  $x_{i_{(p+3)}} = X_{M_{(p+3)}}$  and  $y_{i_{(p+3)}} = Y_{M_{(p+3)}}$ .

6. Search of the fragments using the current ordinate  $y_j$  to a final value  $y_{jF}$ . Achieving the  $y_j$  point by current coordinates along the ordinate axis in the case of  $j = n$ .

7. In the case of  $y_j$  has  $j = n$  there are searching for fragments along abscissa axis. Transition to zero current coordinates on the ordinate axis  $y_j$ , where  $j = 0$  and increment of the current coordinate  $x_i$  along the abscissa axis from the coordinates of the point C of fragment  $L_{(p+2)}$  detected in stage 4, so that  $x_i > x_{i_{(p+2)}}$ .

8. Detection of adjacent fragment  $L_{(p+4)}$  at the point H from the coordinates of the point C in the case of equality of the current coordinate  $x_i = x_{i_{(p+4)}}$  (Fig. 2), to the nearest center of mass  $X_{M_{(p+4)}}$ ,  $Y_{M_{(p+4)}}$  of the point H. A good search for an adjacent fragment  $L_{(p+4)}$  is considered for coordinates  $x_{i_{(p+4)}} = X_{M_{(p+4)}}$  and  $y_{i_{(p+4)}} = Y_{M_{(p+4)}}$  if the condition of belonging to a grouping of adjacent fragments (3) performs for the point C of fragment  $L_{(p+2)}$  and detected point H of fragment  $L_{(p+4)}$ .

9. Search of adjacent fragments with increment along the ordinate axis from the coordinate of already detected point H with coordinates  $X_{M_{(p+4)}}$  and  $Y_{M_{(p+4)}}$ . The condition for a successful search for adjacent fragment of point H is performing of condition (3) for detected point.

10. If  $y_j$  achieve the values along the ordinate axis  $j = n$  that it is necessary to go to stage 2.

11. If  $x_i$  achieve the values along the abscissa axis  $i = m$  that it is necessary to go to stage 2.

12. If for the current coordinates  $x_i$ ,  $y_j$  the condition  $i = m$  and  $j = n$  is accordingly performed and for  $x_{ip}$ ,  $y_{jp}$  the condition  $x_{ip} = X_{MF}$  и  $y_{jp} = Y_{MF}$ , the search for adjacent fragments for the image of the semiconductor wafer can stop.

#### 4. Discussion of the results

The proposed grouping method of image fragments of adjacent dislocation etch pits of the semiconductor wafer detects as fragments falling into adjacent category by the condition (3), and partially adjacent fragments. This is necessary given the characteristics of the binarized image of dislocation etch pits.

Partially adjacent fragments are detected according to the point spaced apart by no more than the proximity measure  $Z$ , according to the conditions (3) and the center of mass of the corresponding fragment by reduction of loop line [14]. The method for determining the loop line width and recovery of loop lines branching are also needed [14].

The method uses the coordinates of the fragments directly into the space of image pixels. The minimum value of the grid size corresponds to one pixel of a particular raster image, however, given the transition by the coordinates of the centers of mass of fragments, the method can be significantly accelerated.

The developed method can have application not being limited by detection of adjacent dislocation etch pits of the semiconductor wafer.

In the absence of reliable signs for grouping neighboring parts of a sustainable cluster with clearly known general purpose, the method allows to combine these disparate fragments based on proximity measure.

Such advantage of the method may be useful in the processing of satellite and other images where it is difficult or impossible to automatically group parts into one clearly known object.

The grouping method of image fragments of adjacent dislocation etch pits of the semiconductor wafer makes it possible to combine neighboring fragments in the group of dislocation etch pits in case of differences of its shape from the standard etch pits.

In the future, this method will be included in the package of methods providing quality monitoring information technology of semiconductor single crystals.

Currently, the authors are working on the implementation of this method.

#### 5. Conclusions

The problem of grouping of loop fragments of dislocation etch pits of gallium arsenide single crystal wafer with orientation [100] after selective etching and polishing is investigated in the article.

The possibility of grouping of adjacent fragments of etch pits are shown.

The grouping method of image fragments of adjacent dislocation etch pits of the semiconductor wafer that determined intermittent dislocations in the binarized image on example of the arsenide single crystal wafer is first proposed.

The method is a part of the quality monitoring information technology of semiconductor single crystals.

The grouping method of image fragments of adjacent dislocation etch pits of the semiconductor wafer has the following advantages:

1. Grouping of significant fragments of binary image of etch pits of dislocation loops allows determining the significant adjacent fragments with uncharacteristic signs for fragments of dislocation etch pits on the semiconductor wafer.

2. Using of the coordinates of the fragments directly into the space of the image pixels without reference to a standard grid of coordinates helps to increase accuracy and reduce intensity of calculations.

3. The possibility to calculate the parameters of the fragments in the image plane with the grid size corresponding to not only the lowest step of image pixel, but also using coordinates of the centers of mass of fragments. This significantly speeds up the grouping of image fragments of adjacent dislocation etch pits of the semiconductor wafer.

### References

- [1] Samoilov, A. N., Shevchenko, I. V. (2013). Metod obnaruzheniya linij konturov v yarkostnykh perepadakh predpolagaemykh granej binarizovannogo izobrazheniya sledov dislokacij na plastinakh GaAs. Avtomatizirovannye sistemy upravleniya i pribory avtomatiki, 165, 22–27.
- [2] Jain, A. K., Murty, M. N., Flynn, P. J. (1999). Data clustering: a review. ACM Computing Surveys, 31, 3. doi: 10.1145/331499.331504
- [3] Zadeh, L. A. (1999). From Computing with Numbers to Computing with Words – From Manipulation of Measurements to Manipulation of Perceptions. IEEE Transactions on Circuits and Systems, 45, 105–119.
- [4] Fung, G. (Ed.) (2001). A Comprehensive Overview of Basic Clustering Algorithms (Glenn Fung). Available at: <http://citeseerx.ist.psu.edu/viewdoc/similar?doi=10.1.1.5.7425&type=ab>
- [5] Du, K.-L. (2010). Clustering: a neural network approach. Neural Networks, 23, 89–107. doi: 10.1016/j.neunet.2009.08.007
- [6] Sheikholeslami, G., Chatterjee, S., Zhang, A. (1998). “WaveCluster: A multiresolution clustering approach for very large spatial databases,” in Proc. 24th VLDB Conf., 428–439.
- [7] Han, J., Kamber, M., Tung, A. K. H.; Miller, H., Han, J. (Eds.) (2001). Spatial clustering methods in data mining: A survey. Geographic Data Mining and Knowledge Discovery, 188–217. doi: 10.4324/9780203468029\_chapter\_8
- [8] Ester, M., Kriegel, H. P., Sander, J., Xu, X.; Simoudis, E., Han, J., Fayyad, U. (Eds.) (1996). A density-based algorithm for discovering clusters in large spatial database with noise. Knowledge Discovery in Databases and Data Mining: Proc. Int. Conf. N. Y.: AAAI Press, 226–331.
- [9] Ankerst, M., Breunig, M. M., Kriegel, H.-P., Sander J. (1999). OPTICS: ordering points to identify the clustering structure. Proc. 1999 ACM SIGMOD Intern. Conf. on Management of data, 49–60.
- [10] Hinneburg, A., Keim, D. A. (1998). An efficient approach to clustering in large multimedia databases with noise. Proc 4th Intern. Conf. on Knowledge Discovery and Data Mining. N. Y., 58–65.
- [11] Hinneburg, A., Gabriel H.-H. (2007). DENCLUE 2.0: Fast clustering based on kernel density estimation. In Proc. 2007 Int. Conf. Intelligent Data Analysis (IDA'07), 70–80. doi: 10.1007/978-3-540-74825-0\_7
- [12] Samoilov, A., Shevchenko, I. (2016). Development of methods for separation of binarized fragments of etching pits of semiconductor wafer. Technology Audit And Production Reserves, 3 (1(29)), 60–68. doi: 10.15587/2312-8372.2016.71988
- [13] Samoilov, A. N. (2013). Analiz adaptivnoy porogovoy obrabotki yarkostnykh perepadov elementov cifrovogo rastrovogo izobrazheniya. Materialy IX Mizhnarodnoi naukovo-praktichnoi konferencii «Rozvitok naukovich doslidzhen 2013», 26–28.
- [14] Samoilov, A. N. (2015). Methods for recovering the dislocations contour line of gallium arsenide wafer of digital image. Eastern-European Journal of Enterprise Technologies, 3 (5 (75)), 8–16. doi: 10.15587/1729-4061.2015.43326